2 FIA

PREDICTING THE PROBABILITY OF STAND DISTURBANCE[1]

Gregory A. Reams and Joseph M. McCollum[2]

Abstract—Forest managers are often interested in identifying and scheduling future stand treatment opportunities. One of the greatest management opportunities is presented following major stand level disturbances that result from natural or anthropogenic forces. Remeasurement data from the Forest Inventory and Analysis (FIA) permanent plot system are used to fit a set of models that predict the probability of harvest for the five FIA survey units of Georgia. We assume a logistic function that establishes asymptotes at 0 and 1. We found that geographic region, ownership, number of trees per acre, and average stand diameter are correlated to the probability of harvest. A plot was considered harvested if any tree > 5 in. d.b.h. was cut. The average probability of harvest over the last 8.5 years was approximately 33 percent for the central and southern regions of Georgia. The average rate of harvesting in northern Georgia was 21 percent. These models can be used to predict the probability of harvest for a set of stand conditions and when combined with area expansion factors, to estimate the acreage of harvested stands.

INTRODUCTION

A study to model disturbance rates resulted from two potential uses. Firstly, the recent move to annual forest inventories by the Forest Inventory and Analysis (FIA) program of the USDA Forest Service, Southern Research Station has raised some interesting questions regarding efficient allocation of sample plots. Specifically, the greatest change in an inventory often occurs in conjunction with major disturbances. A nonexhaustive list of disturbances that can significantly alter an inventory includes harvesting, insects and disease, and wind damage and other weather related events.

In the Southern U.S., harvesting is the largest disturbance and thus, estimating the probability of harvest and comparing the actual rates of disturbance to disturbance detection techniques is a prerequisite for implementation of an inventory based on disturbance detection. Second, and more important to the silvicultural community, is the need to estimate the type and acreage of stands available for site preparation, planting, or other treatment opportunities following harvest. The recently completed seventh forest survey of Georgia (Thompson 1998) provides an ideal opportunity to model harvesting rates.

INVENTORY METHODS

Estimates of change and rates of change were available from the 1997 remeasurement of 5,386 permanent sample plots established in the previous survey in 1989. The plot design for the previous inventory was based on a cluster of 10 points. Variable radius plots were systematically spaced within a single forest condition at three to five points. At each point, trees ≥5.0 in. d.b.h. were selected for measurement on a variable radius plot defined by a 37.5-factor prism. Trees < 5.0 in. d.b.h. were tallied on a fixed-radius plot around each plot center.

MODEL DEVELOPMENT

When the dependent variable is an indicator variable, the shape of the response function will often be curvilinear. For our specific application of modeling harvest rates, the response variable is either 1 or 0 depending on whether an inventory plot was harvested or not. An additional necessary requirement is to use a model that has asymptotes at 0 and 1. The logistic function meets all of these prerequisites and has been used to model event probabilities such as individual tree mortality for several decades (Monserud 1976). The general form of the logistic function is:

$$p(h) = \frac{\exp(X)}{1+\exp(X)} + \varepsilon \tag{1}$$

where

$p(h)$ = the probability of harvest,
$\exp(X)$ = the exponential function e^x,
X = the set of predictor variables, and
ϵ = the error term.

The probability of harvest is related to many variables, including volume per acre by tree species, size, quantity, and quality of the trees, and the species mix per unit area. The correlation between these potential predictor variables at time 1 (1989 measurement) and harvesting rates at time 2 (1997 measurement) can be easily investigated with FIA's remeasurement data. Other potentially influential variables such as operability, distance to mill, distance to roads or urban populations (Wear and others 1999), and planted versus natural stands are variables that were not included in this study. With the exception of planted versus natural stands, these variables could not be investigated because the variables are either not collected or readily available for analysis. Future studies are planned to investigate the excepted variables.

RESULTS AND DISCUSSION

The logistic regression model (1) was fit to the 1997 FIA remeasurement data for the five survey units (Southeast, Southwest, Central, North Central, and North) of Georgia (fig. 1). A sample plot is considered harvested if any tree (> 5 in. d.b.h. at time 1) was cut. This study does not distinguish levels of cutting and includes the full spectrum of harvesting practices. The average remeasurement interval across the state was 8.5 years. The average tree diameter by species group and number of trees per acre by species group are listed for each of the survey units (table 1).

The fitted coefficients to model (1) for each of the survey regions are listed in table 2. Under each survey unit are

[1] Paper presented at the Tenth Biennial Southern Silvicultural Research Conference, Shreveport, LA, February 16-18, 1999.

[2] Supervisory Mathematical Statistician and Mathematical Statistician, USDA Forest Service, Southern Research Station, Forest Inventory and Analysis, Asheville, NC 28802, respectively.

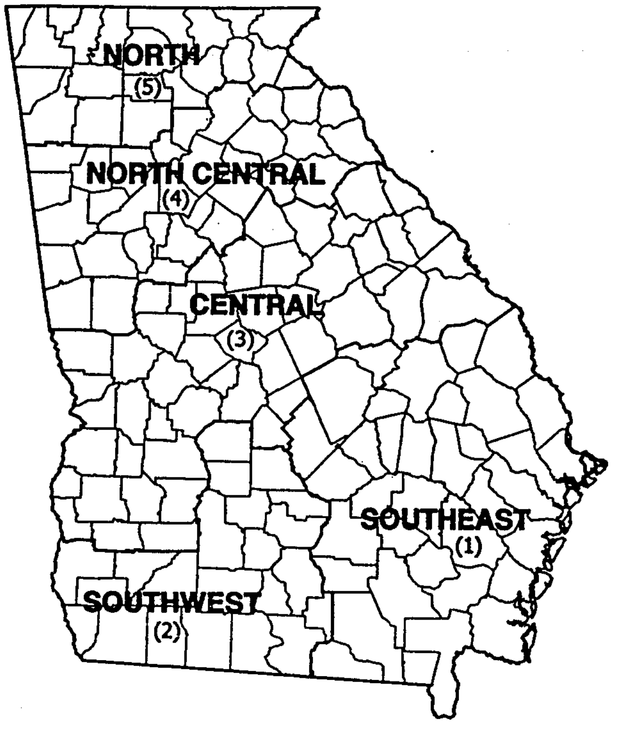

Figure 1—The five forest survey regions in Georgia.

forward selection procedure for variable inclusion. Because users desire to differentiate harvest event probabilities by ownership, we retained all ownership variables in the models. With the exception of forest industry (FORCO) lands, nonindustrial private forest (NIPF) lands , and other corporate (CORP) lands, for both the interaction and main effects model in the Northern survey unit and the interaction model in the North Central survey unit, all ownership variables are different from each other at the p = 0.05 level. Forest industry, private nonindustrial, and other corporate lands have harvest rates greater than the government but are not distinguishable from each other.

Other significant variables (p = 0.05) in the models include mean initial stand diameter (d.b.h.), mean initial pine stand diameter (PDBH), mean initial hardwood stand diameter (HDBH), number of initial pine trees per acre (PTPA), number of initial hardwood trees acre (HTPA), and a number of interaction terms (table 1). All mean stand diameter and trees per acre variables are based on sample plots with trees > 5 in. d.b.h. Plots with no trees > 5 in. d.b.h. were excluded from model fitting.

Interpretation of the model coefficients proceeds as follows. Positive coefficients are associated with increased rates of harvest and negative coefficients with decreased rates of harvest. For example, in the main effects model for survey unit 1, other corporate lands are harvested at higher rates than all other ownership categories (table 2). The variables that are the most difficult to interpret are the interaction terms between an ownership category and a stand level variable, such as PDBH. For example, for Southeastern (unit 1) Georgia there is an important positive interaction between NIPF lands and PDBH. The interpretation of this interaction is that as average pine diameter increases harvest rates accelerate faster on NIPF lands than on other land ownerships.

Table 1—Average tree diameter and numbers of trees per acre for trees >5 inches d.b.h. by species group (all species, hardwoods, and conifers). Plots with no trees >5 inches d.b.h. were excluded from the analysis. Average d.b.h. and number of trees per acre are based on time 1 measurements (1989). The Georgia survey units are the Southeastern (unit 1), Southwestern (unit 2), Central (unit 3), North Central (unit 4), and Northern (unit 5)

	Survey unit				
	1	2	3	4	5
Mean d.b.h.	11.37	12.18	11.63	11.59	11.18
Mean pine d.b.h.	10.30	11.70	11.26	10.68	11.10
Mean hardwood d.b.h.	12.68	12.69	11.93	12.31	12.30
No. trees/ac	181.4	154.8	152.7	164.1	170.6
No. pine trees/ac	109.3	79.1	68.6	77.2	62.1
No. hardwood trees/ac	72.1	75.7	84.1	86.9	108.5

two models, the first is a main effects model and the second, an interaction effects model. The reason two different models are listed is to illustrate that a main effects model is overly restrictive and does not allow for the non-linear relationship between the predictor variables and predictant (fig. 2 and 3).

We used the logistic regression procedure in Statistical Analysis System (SAS) to fit the models and used the

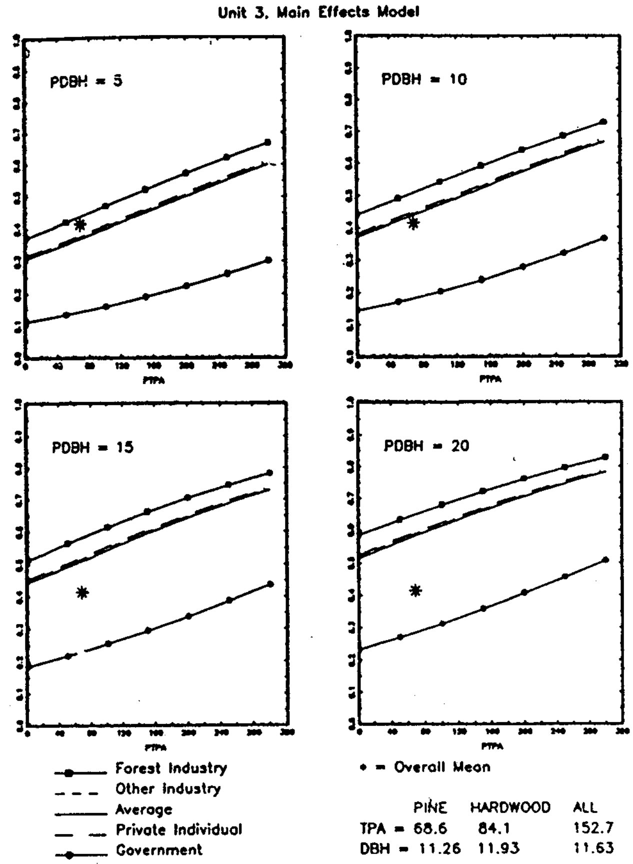

Figure 2—The conditional probability of harvest as given by the main effects model for survey unit 3 (Central) Georgia, by average stand pine diameter (trees > 5 in. d.b.h.), number of pine trees/acre (trees > 5 in. d.b.h.), and ownership class. The lower right corner lists the mean sample value for all continuous variables in the model. The asterisk represents the average conditional probability of harvest for the survey unit.

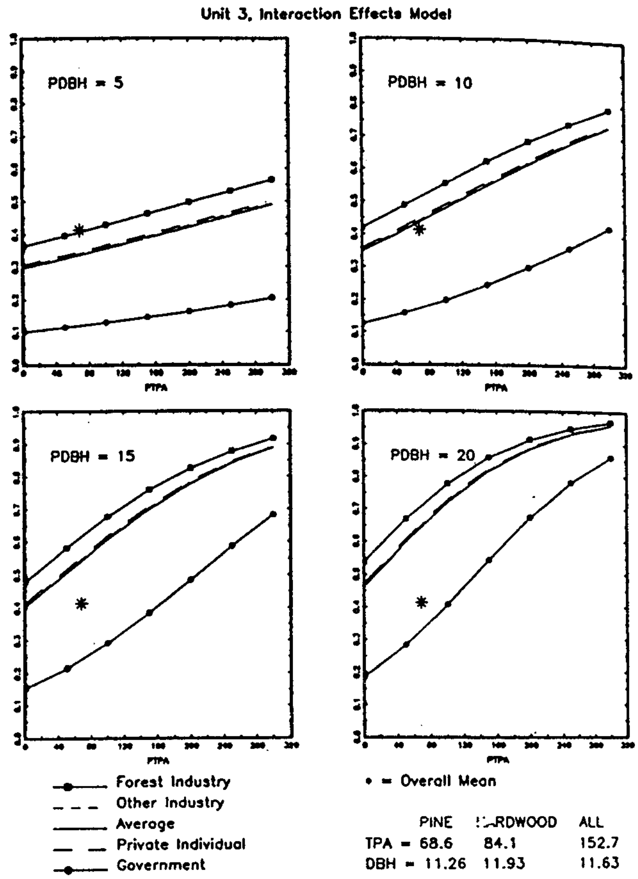

Figure 3—The conditional probability of harvest as estimated from the interaction effects model for survey unit 3 (Central) Georgia, by average stand pine diameter (trees > 5 in. d.b.h.), number of pine trees/acre (trees > 5 in. d.b.h.), and ownership class. The lower right corner lists the mean sample value for all continuous variables in the model. The asterisk represents the average conditional probability of harvest for the survey unit.

In general, across the five survey units in Georgia, harvest probabilities are greatest on sites with large pines and many of them. Ownership harvest rates vary by region; however, rates are always the lowest government controlled lands. Although there is the perception that harvest rates are greatest on FORCO lands, we found that this is not always true. For example, in southeastern (unit 1) Georgia, harvest rates are greatest on CORP lands and NIPF lands. Overall across the State, FORCO lands are harvested at the greatest rate, however NIPF and CORP lands have harvest rates approaching those of FORCO, and sometimes exceed FORCO especially in southern Georgia.

Harvest events over the last 8.5 years in Georgia have varied by region, with the greatest rates in central and southern Georgia. The rates gradually decrease moving to the northern region. The conditional probability of a harvest event (conditional on including only plots with trees >5 in. d.b.h.) in survey units 1 and 2 is 40 percent, 43 percent in unit 3, 38 percent in unit 4 and 23 percent in unit 5. Harvest event rates appear at first glance as extremely high, but remember this includes all types of harvesting, and we conditioned the data set to include only plots with trees >5 in. d.b.h. at time 1. The chance of a harvest event based on all plots is 31 percent for unit 1, 32 percent for unit 2, 35 percent for units 3 and 4, and 21 percent for unit 5. Final harvest occurred on 16 percent of all forestland in Georgia over the 8.5-year period (Thompson 1998). Final harvest is defined as the removal of the majority of the merchantable trees in a stand, leaving residual stand stocking less than 50 percent.

Example Using the Interaction Model for Unit 3

Assume the probability of harvest for a stand with the following attributes is needed. Assume the stand is on FORCO lands, and has a mean stand pine diameter (trees >5 in. d.b.h.) of 10 in. d.b.h., and 200 pine trees per acre (trees >5 in. d.b.h.). The interaction model coefficients for unit 3 (table 2) are:

$$X = -2.4223 + 1.6240(FORCO) + 1.3623(NIPF) + 1.3354(CORP) + 0.0473(PDBH) + 0.00055(PDBH)(PTPA).$$

Table 2—Estimated coefficients for the probability of harvest models for the five FIA survey units in Georgia. Insert the value of X into equation (1) to predict the probability of harvest. n_1 = number of plots harvested, n_2 = number of plots not harvested, $n = n_1+n_2$

Unit 1 n = 1636, n_1=653, n_2 = 983

Main effects model

X = -2.5436+1.4353*(FORCO)+1.6621*(NIPF)+1.9429*(CORP)
+0.0331*(PDBH)-0.0014*(HDBH)+0.0039*(PTPA)

Interaction effects model

X = -2.2400+1.4488*(FORCO)+1.5454*(NIPF)+1.9043*(CORP)
+0.0366*(NIPF)*(PDBH)-0.0026*(NIPF)*(PTPA)+0.0005*(PDBH)*(PTPA)
-0.0002*(HDBH)*(HTPA)

Unit 2 n=709, n_1=274, n_2=435

Main effects model

X = -2.3470+1.6563*(FORCO)+1.2819*(NIPF)+1.4056*(CORP)
+0.0567(PDBH)+0.0027(PTPA)-0.0033(HTPA)

Interaction effects model

X = -2.0289+1.5842*(FORCO)+0.6564*(NIPF)+1.4990*(CORP)
+0.0554*(NIPF)*(PDBH)+0.0005*(PDBH)*(PTPA)-0.00004*(PTPA)*(HTPA)

Unit 3 n=1307, n_1=540, n_2=767

Main effects model

X = -2.3601+1.5381*(FORCO)+1.2954*(NIPF)+1.2721*(CORP)
+0.0581*(PDBH)+0.00041*(PTPA)

Interaction effects model

X = -2.4223+1.6240*(FORCO)+1.3623(NIPF)+1.3354(CORP)
+0.0473(PDBH)+0.00055*(PDBH)*(PTPA)

Unit 4 n=736, n_1=279, n_2=457

Main effects model

X = -1.5705+0.5910*(FORCO)+0.0577*(NIPF)+0.4887*(CORP)
+0.0695*(PDBH)+0.0040*(PTPA)

Interaction effects model

X = -1.1042-2.9422*(FORCO)+0.0731*(NIPF)+0.4979*(CORP)
+0.2749*(FORCO)*(DBH)+0.1446*(FORCO)*(PDBH)+0.0497*(PDBH)
-0.0353*(HDBH)+0.0004*(PDBH)*(PTPA)

Unit 5 n=549, n_1=126, n_2=423

Main effects model

X = -1.4292+0.2123*(FORCO)+0.1699*(NIPF)+0.3415*(CORP)
+0.0476*(PDBH)+0.0019*(PTPA)-0.0044*(HTPA)

Interaction effects model

X = -1.1038+0.2054*(FORCO)+0.1084*(NIPF)+0.2945*(CORP)
+0.0514*(PDBH)-0.00055*(HDBH)*(HTPA)

To obtain the predicted probability proceed as follows:

$$X = -2.4223+1.6240(1)+1.3623(0)+1.3354(0)+0.0473(10)+0.00055(10)(200)$$
$$X = 0.7746$$

Inserting X = 0.7746 into model (1) results in:

$$p(h) = \frac{\exp(0.7746)}{1+\exp(0.7746)} = 0.6845$$

The predicted probability of harvest for a similar stand on government controlled lands is calculated as follows:

$$X = -2.4223+1.6240(0)+1.3623(0)+1.3354(0)+0.0473(10)+0.00055(10)(200)$$
$$X = -0.8494$$

Inserting X = -0.8494 into model (1) results in:

$$p(h) = \frac{\exp(-0.8494)}{1+\exp(-0.8494)} = 0.2995$$

CONCLUSION

We have demonstrated the utility of FIA data for harvest event rate modeling and found that region, ownership, average tree size, and the number of trees per acre influence rates. Rates of harvest increase with tree size and numbers of trees per acre, with a greater preference to pine. The greatest rates of harvest are in central and southern Georgia. We recommend the use of a logistic regression model or similar models where the response variable is bounded by zero and one. Models should include possible interaction terms and not restrict parameter estimation to main effects.

The number of acres impacted by harvest can be estimated by assigning a harvest probability to each plot and multiplying by an appropriate plot expansion factor and summing the resulting product. Future research will address final harvest models and, in addition to the variables investigated in this study, evaluate the possible effects of planted versus natural stands, accessibility, operability, and tract size. Thompson (1999) found that tract size is correlated to removal rates on NIPF lands in Florida.

Users of the harvest event models must be mindful that all models are based on plots with at least one tree greater than 5 in. d.b.h. Expansion of the modeled rates to an area basis must adjust for the conditional plot selection used in model development.

ACKNOWLEDGMENTS

The authors thank Mike Thompson, Callie Jo Schweitzer and Ray Sheffield for their insightful review comments and advice.

REFERENCES

Monserud, R.A. 1976. Simulation of forest tree mortality. Forest Science. 22: 438-444.

Thompson, M.T. 1998. Forest statistics for Georgia, 1997. Resour. Bull. SRS-36. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 92 p.

Thompson, M.T. 1999. A forested tract-size profile of Florida's NIPF landowners. Res. Pap. SRS-15. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 10 p.

Wear, D.N.; Liu, R.; Foreman, J.M. [and others]. [In press]. The effects of population growth on timber management and inventories in Virginia. Forest Ecology and Management.